Data entry and data accuracy

Hartzell V. Schaff, MD,[a] Morgan L. Brown, MD, PhD,[b] and Judy R. Lenoch[a]

Hamlet: Madam, how like you this play?
Queen Gertrude: The lady doth protest too much, methinks.
Hamlet. Act 3, scene 2

The critique of our article by Dr Shahian and associates can be separated into three areas.[1,2] First, they raise issues with methodology and question use of a previous database version as well as the selection of study variables. We purposely chose to abstract older records so that the fellows and residents who participated in the study would have no firsthand knowledge of the patients and their procedures. These data were then compared with information in the database (version 2.52.1), so it was necessary to use that version of the database for the new abstraction. The fact that version 2.52.1 was phased out in December 2007 seems irrelevant to our findings inasmuch as all of the data collected with that version of the database remain in place and continue to be used for quality assurance and outcome studies. The decision to study variables linked to valve disease simply relates to the fact that this was the largest group of patients entered into our database at the time of the study and continues so in current practice. Thus, it is important information to us.

On the use of untrained data abstractors, we agree with the authors and believe the data accuracy and completeness are optimized when records are abstracted by dedicated, trained personnel. But Shahian and associates answer their own question of why coding performance of untrained data abstractors should be studied when, 11 paragraphs later, they advocate entry of operative variables by the surgeon. It seems inevitable that in many hospitals this task will be delegated to assistants, which, in most academic institutions are residents and fellows. Indeed, an important question is, who enters data in the current Society of Thoracic Surgeons Adult Cardiac Database (STS ACD)? Do all centers have dedicated coders or do clinical or secretarial personnel contribute to data entry? What model results in the most complete and accurate information? We suspect that busy personnel whose primary responsibilities are clinical or administrative will be less thorough in record review than dedicated coders, and this issue seems worthy of study by the database committee.

After acknowledging that classification of valve etiologic anatomy, pathophysiology, and procedures have been incomplete in the past, Shahian and associates write that the upcoming database version 2.70 will have an expanded section on valve disease. Any changes in the STS ACD that result in greater precision and clarity are welcome, but the proposed changes may have the opposite effect because of the increased burden of data entry that will be required. In its current version, 2.61, there are approximately 335 data fields. The proposal for version 2.70 circulated in August 2010 adds 289 fields (306 new fields, 40 deleted fields, and 17 deleted and replaced fields). This extra work is not necessarily offset by simplification of data entry, inasmuch as much of the additional effort is in review of the medical record. Surgeons, data managers, and supporting staff may question the need for additional information that appears driven by research interests for a database that has as its central purpose quality assurance. For comparison with the proposed version 2.70 STS ACD, which has 624 fields, the Northern New England Cardiovascular Disease Study Group fulfills its quality assurance responsibility with a database that contains 203 fields (version 6.0), and the New York State database lists 137 fields (New York State Cardiac Advisory Committee, 7/09). The size and complexity of the STS ACD has important implications as regards cost as well as data accuracy, and this is not a trivial matter. If the database is expanded as proposed, many centers will not be able to keep up with data entry and verification with current staffing, compounding the financial strain of the ''unfunded mandate'' of quality assurance.

The balance of the article discusses previous analyses and our reference to the study of Herbert and colleagues.[3] Shahian and coworkers believe we have misrepresented the findings of this investigation, and although we may differ in interpretation, there should be no confusion about what was written. The first two sentences from the Results section of that article read: ''Complete correlation of the clinical record and the database occurred for 5 (2.0%) of the patients. An additional 190 records (82.2%) had one to ten discrepancies, and 39 (16.8%) had more than ten fields of disagreement.'' Readers of the *Journal* are encouraged to review the article and make their own conclusion.

Finally, the authors seem to suggest that, like financial institutions that are too big to fail, the STS ACD is too large and valuable to examine critically. This attitude is puzzling

From the Mayo Clinic,[a] Rochester, Minn; and University of Alberta,[b] Edmonton, Alberta, Canada.

Address for reprints: Hartzell V. Schaff, MD, Division of Cardiovascular Surgery, Mayo Clinic, 200 First St SW, Rochester, MN 55905 (E-mail: schaff@mayo.edu).

J Thorac Cardiovasc Surg 2010;140:960-1
0022-5223/$36.00

doi:10.1016/j.jtcvs.2010.09.006

because our study of variability in data abstraction highlights an area where process improvement will strengthen the database and information derived from it. The conclusion from our investigation is straightforward: data variability increases when abstraction and entry are performed by clinical staff rather than dedicated database personnel. We welcome the opportunity to review evidence to the contrary.

References

1. Shahian DM, Edwards F, Grover F, Jacobs JP, Wright CD, Prager RL, et al. The Society of Thoracic Surgeons National Adult Cardiac Database: a continuing commitment to excellence. *J Thorac Cardiovasc Surg*. 2010;140:955-9.
2. Brown ML, Lenoch JR, Schaff HV. Variability in data: the Society of Thoracic Surgeons National Adult Cardiac Surgery Database. *J Thorac Cardiovasc Surg*. 2010;140:267-73.
3. Herbert MA, Prince SL, Williams JL, Magee MJ, Mack MJ. Are unaudited records from an outcomes registry database accurate? *Ann Thorac Surg*. 2004; 77:1960-4.